

Western University
Scholarship@Western

Department of Economics Research Reports

Economics Working Papers Archive

1981

On the Short Run Demand for Money and the Real Balance Effect

David Laidler

Follow this and additional works at: <https://ir.lib.uwo.ca/economicsresrpt>



Part of the [Economics Commons](#)

Citation of this paper:

Laidler, David. "On the Short Run Demand for Money and the Real Balance Effect." Department of Economics Research Reports, 8105. London, ON: Department of Economics, University of Western Ontario (1981).

20096

ISSN: 0318-725X

ISBN: 0-7714-0264-3

RESEARCH REPORT 8105
ON THE SHORT RUN DEMAND FOR MONEY AND
THE REAL BALANCE EFFECT
by
David Laidler

April 1981

Department of Economics Library

APR 1981

University of Western Ontario

ON THE SHORT RUN DEMAND FOR MONEY AND
THE REAL BALANCE EFFECT*

by

David Laidler

* I have benefitted greatly from correspondence with Dr. William White of the IMF in developing the argument of this paper, while those who are familiar with the work of Peter Jonson will recognise how much I owe to discussions with him. My colleagues, Joel Fried and Peter Howitt, have made important and helpful comments on an earlier draft of this essay. Nevertheless, the arguments presented here, right or wrong, are my own responsibility. This paper has been written while on Sabbatical leave from The University of Western Ontario, to which Institution, along with the Social Science and Humanities Research Council of Canada, I am grateful for financial support.

I

THE LONG AND SHORT RUN DEMAND FOR MONEY

No proposition in macro-economics has received more attention than that there exists, at the level of the aggregate economy, a stable demand for money function. When we say that the demand for money function is stable we mean at the very least that money holdings, as observed in the real world, can be explained, to conventionally acceptable levels of statistical significance, by functional relationships that include a relatively small number of arguments. We also mean, or should mean, that the same equation is capable of being fitted to samples of data drawn from different times and places without it being necessary to change the arguments of the relationship in order to achieve satisfactory results, and also without the estimated quantitative values of the parameters changing too much.

Now of course a cavalier treatment of the requirements that the number of parameters be "relatively small", and that parameter values not change "too much" when the data are changed, could permit claims to be made on behalf of the stability of almost any relationship, but, in this case, there is no need to abuse the English language. In practise a "small" number of arguments has meant three or four - typically including a scale variable such as income or wealth, an opportunity cost variable such as a nominal interest rate or some measure of the expected inflation rate, and, if nominal balances have been the dependent variable, the general price level. The requirement that parameters not change "too much" has meant not only that they have been expected to take their theoretically predicted sign, but also to stay within reasonable quantitative ranges as well, in the region of 0.5 - 1.0 or a little greater for the income elasticity of demand for money, somewhere

around -0.5 or less for the interest elasticity depending upon the interest rate utilised, and since economic theory predicts that the demand for money is a demand for "real" balances, a price level elasticity of demand for nominal balances of close to 1.0 .

I have surveyed empirical work on the demand for money elsewhere, (Laidler 1976, 1980) and there is no need to go into details about these matters again here. It will suffice to note that, although not every test on every set of data has proved satisfactory, the demand for money has turned out to be "stable" in the sense in which I have used the word above quite often enough to convince the majority of monetary economists, some of whom twenty years ago were quite skeptical (see e.g., Modigliani 1978), of the importance of the relationship for our understanding of macroeconomic phenomena. One characteristic of the results generated by empirical work on the demand for money that seems to be rather general is that, unless one is dealing with data which are highly aggregated over time - business cycle phase averages for example - or which cover such a long period of time - fifty years or more say - that the variation in the data used is dominated by secular changes, it has proved necessary to distinguish between the "long run" and the "short run" demand for money in order to achieve satisfactory results.

The long run aggregate demand for money function may be thought of as being generated by the outcome of the following thought experiment: consider an economy in which, given the values of the various arguments in the function, the aggregate of agents will desire to hold a certain quantity of money and are able to do so; then ask how much money they would be observed to hold in various alternative circumstances in which the variables that determine money holding took different values. If the tastes of agents vis a vis money

holding did not change over time, and if there were no barriers to their moving instantaneously from holding one amount of money to another, the outcome of this experiment would be the same as that generated by varying the values of the arguments of the demand for money function over time and then observing the changes in cash balances associated with those variations. However, if agents face any costs of adjusting their money holdings, so the argument goes, they will not move immediately to a new point on their long run demand for money function when one of its arguments changes. They will begin to move towards that point, but the speed of the approach will be determined by their response to the adjustment costs involved in getting there.

It will help with the clarity of the argument to put all this into a familiar algebraic form at this point. It is convenient to work with a log. linear (constant elasticity) form of the demand for money function, and to divide time up into discrete periods, Hicksian "weeks" if you like. Thus, with X a vector of factors determining the demand for real balances, p the log. of the general price level, m^* the log. of the quantity of money demanded as determined by the long run function, we write

$$m^* = f(X) + p \quad (1)$$

which is the long run aggregate demand for money function.

We also write, with the subscript -1 denoting a one period time lag and with m the log. of nominal money demanded

$$m - m_{-1} = b(m^* - m_{-1}) \quad 0 < b \leq 1 \quad (2)$$

which tells us how much of the gap between actual and ultimately desired money holdings will be closed during one period. Together these equations yield the short run demand for money function

$$m = b \left[f(X) + p \right] + (1 - b)m_{-1} \quad (3)$$

An equation such as (3), or something very like it, has been used in an enormous number of studies of the demand for money, and the lagged dependent variable has almost invariably proved an important addition as far as increasing the relationship's explanatory power is concerned. The addition in question has usually been regarded as quite innocuous, for the kind of adjustment cost argument which I have just sketched out has been widely applied, not just to the demand for money, but to the consumption function as well, not to mention the demand for various durable goods (see Harberger 1960). Indeed, the seminal article on this particular aspect of the demand for money, Chow (1966) was explicitly an application to money of techniques that he had used when working on the demand for automobiles, techniques that, in essence, give econometric content to the Marshallian distinction between the long run and short run response of quantity demanded of some good to a change in some argument of its demand function, a distinction that seems at first sight to be universally applicable.

The bulk of this paper will be devoted to elaborating upon the proposition that, notwithstanding the widespread practise of adding a lagged dependent variable to the demand for money function, the belief implicit in this practise that the demand for money in the aggregate economy can be modelled in the long and short runs "as if" money was a consumer durable good, is

fallacious. It will also examine some of the implications of this for what we do and do not know about the aggregate demand for money. The phrase, "in the aggregate economy" is italicised, because the problem to be discussed arises at the level of what Patinkin (1965) called the market experiment, and not at the level of the individual experiment at all. It will nevertheless be helpful to begin the argument with an examination of the relevant individual experiment.

II

THE INDIVIDUAL EXPERIMENT AND THE REAL BALANCE EFFECT

Though it was a controversial matter at one time, it is by now as near to universally accepted as anything in economics ever is that once the relevant object of choice is recognised to be "real balances", money holdings measured in constant-purchasing power terms, and once their durability is taken account of, the individual agent's demand for money can, and indeed should, be analysed along with his demand for everything else. The amount of real balances that he will hold will be the outcome of the interaction of his utility function with his budget constraint, just like the quantity of anything else he will demand. It was at one time widely questioned whether it made sense to argue that money yielded "utility" to the individual in the same way as other goods, but it is by now well established that, from the point of view of the individual experiment, the "story" that one tells about this matter makes no critical difference.

One may argue that, by holding real balances, the individual agent is able to avoid the embarrassment of being unable to pay up promptly when

unexpectedly called upon to meet his obligations in cash (Patinkin 1965); one may argue that he is enabled to cut down on the transactions costs involved in liquidating income earning assets when cash is required (Baumol 1952, Tobin 1956); one may argue that the agent can avoid the uncertainties about future command over resources inherent in holding variable capital value assets such as bonds (Keynes 1936, Tobin 1958); and so on. What is postulated here turns out to make no difference to our ability to integrate the analysis of the demand for money with that of other aspects of the agent's choices, any more than the motives we attribute to the owners of automobiles make any difference to our ability to apply choice theory to analysing the demand for that durable good.

What is important, as Patinkin (1965) chs. 5-7 showed quite clearly is that agents do desire to hold real balances, and not why. This is not to deny that if we formulate some precise hypothesis about the nature and source of the "utility" that money holding yields the individual, we might thereby put ourselves in a position also to formulate more precise hypotheses about the quantitative nature of the agent's demand for money function than generalised choice theory would yield. After all the Baumol-Tobin "square root rule" is an example of just this possibility working out in practice. However, if the qualitative predictions yielded by the basic theory of choice are sufficient for any particular purpose, then there is no need to ask why cash balances yield utility before applying that theory to analysing the demand for money. This is not to say that, when we engage in economic analysis, we never need to pay attention to those special characteristics of money as a social institution that facilitates the processes of exchange, but it is to say that, when considering the money holding

behaviour of an individual agent acting alone, it will suffice to treat real money balances "as if" they are a service-yielding consumer durable.

Using the same symbols as before, but attaching to them, where appropriate, the subscript i to indicate that we are indeed dealing with an individual, we may write the individual agent's long run demand for money function, where "long run" is defined in a manner analogous to that already used above, as

$$m_i^* = f_i(X)_i + p \quad (4)$$

Here, it is as well to note explicitly that m_i^* refers to the quantity of nominal balances the individual will plan to end up holding at the end of the current period, given the price level and the values of the variables included in X , that rule during the period, if he faces no costs of adjusting his money holdings. If we are willing to entertain the possibility of our agent being off his long run demand for money function over a time span longer than one period, and we should be, if only because Archibald and Lipsey (1958) established one set of mechanisms that make this a reasonable postulate, we might also argue that he moves back towards it slowly according to

$$(m_i - m_{i-1}) = b(m_i^* - m_{i-1}) \quad (5)$$

Once again then, we can derive a short run demand for money function of the conventional form

$$m_i = b \left[f_i(X)_i + p \right] + (1 - b)m_{i-1} \quad (6)$$

Here m_1 is the amount of money the agent actually plans to hold at the end of the current period, and m_{1-1} is the amount of money with which he begins the period. This, as we shall see in a moment, may or may not be the amount of money he chose to hold in period -1. Be that as it may, the adjustment parameter b is easily enough motivated in the individual experiment. If, by the end of the period for which he is choosing his cash holdings, our agent is not on his long run demand for money function, he obviously enjoys less utility, or incurs greater costs somewhere or other, than he otherwise would. Suppose, however, that in moving back towards that long run relationship over that period he also incurs transactions costs of some sort. Call the first cost K_1 and the second K_2 , and let them be determined in the following way.

$$K_1 = \alpha_1 (m_1^* - m_1)^2 \quad (7)$$

$$K_2 = \alpha_2 (m_1 - m_{1-1})^2 \quad (8)$$

The agent seeking to minimise the sum of these costs will adjust his cash balances over time according to equation (5) where

$$b \equiv \frac{\alpha_1}{\alpha_1 + \alpha_2} \quad (9)$$

The above cost functions are undoubtedly arbitrary; their quadratic form has much more to do with the fact that this enables us to derive a linear, and therefore easy to handle, adjustment process, than with any well

thought through microeconomic analysis; also, the existence of an adjustment cost function such as (8) for the individual is not easily reconciled with the lump sum adjustment costs that are sometimes used in deriving the long run demand for money function (when, for example, the Baumol-Tobin inventory approach is used), but such criticisms are not of any great importance for present purposes. Nothing fundamental in the arguments that follow depends upon the linearity of the adjustment process, but the simplicity of the argument is enhanced if we make the arbitrary assumptions that have to be made in order to keep the individual's short run demand for money function in the form given by (6), not least because that is the form which is usually thought of as underlying the similar relationship used in empirical work on aggregate data.

Equation (6) tells us that the amount of nominal money we will observe our individual holding at the end of any time period will depend upon the general price level at which trade takes place during the period, whatever factors we might put in X , let us say real income and a representative nominal interest rate over the period, and the quantity of nominal money he held at the beginning of that period. Such an equation makes perfectly good economic (and econometric) sense. The price level and nominal interest rates are quite beyond the individual's control, as is real income - unless we go into a model in which the labour leisure choice is endogenous, and if we did that we ought to put the real wage, and some endowment of labour power into the relationship instead. Beginning of period money is also exogenous from the point of view of current period behaviour, however it may be determined. The only endogenous variable in the equation is indeed the one that appears on its left hand side.

Equation (6) is a meaningful, if rather trivial expression, that can form the basis for a series of equally meaningful, and equally trivial, individual experiments. We can start our individual out on his long run demand for money function, face him with changes in his real income, the interest rate, or the price level, and use equation (6) to generate the resulting time path of his nominal money holdings. In this case, note that m_{-1} , beginning of period money holdings, will be given by the value taken by the dependent variable of equation (6) in period -1. We can also present our individual with a windfall gain in nominal money holdings (perhaps as a result of the passage of a helicopter, c.f., Friedman (1969)) hold interest rates, prices, and his income, constant, and once again use equation (6) to tell us about his reaction. In this case, of course, beginning period money will not be equal to the individual's money holdings at the end of period -1.

We do not usually come across this last experiment in the discussions of the demand for money, finding it instead in discussions of the "real balance effect" where the influence of money on expenditure flows is at the centre of attention. However, it is a point too often taken for granted, and perhaps for that reason not fully enough appreciated, that, in the individual experiment, whenever there arises a discrepancy between desired long run money holdings and actual money holdings, for no matter what reason, there must be accompanying effects on expenditure flows, either on current consumption goods and/or on the acquisition of other assets. The change in money holdings on the left hand side of equation (5) must have its counterpart in the agent's expenditure if his budget constraint is to be satisfied. That is to say, when we talk of the adjustment over time of the agent's money holdings towards their long-run equilibrium we are also talking about what is, to all intents

and purposes, a real balance effect. This is true whether the experiment we are describing is set in motion by a variation in the arguments of the agent's long run demand for money function, or by a change in his endowment of nominal money.

I am here using the phrase "real balance effect" in a rather broader sense than did Patinkin (1956), because he reserved the term to characterise only wealth effects. He used the phrase "substitution effect" to describe the consequences of those disturbances to the individual that required him to change only the composition of his assets. Here I am bringing both types of reaction under the one heading. Moreover, this analysis of the dynamics whereby desired long run money holdings are reached presented above is different from the account offered by Archibald and Lipsey (1958), to which I have already alluded, in their extension to the multi-period case of Patinkin's original analysis of the operation of the real balance effect in the individual experiment. However, these differences reflect the fact that the foregoing analysis has started from the literature on the Demand for Money, rather than that which deals with the integration of Monetary Theory and General Equilibrium Theory. They do not imply that the conclusions that have been stated above about the relationship between monetary adjustments and expenditure flows are in any way misplaced. Thus, we may use the insights yielded by Patinkin's and Archibald and Lipsey's work to illuminate the connection between the individual and market experiments in the analysis of the role of adjustment costs in the demand for money function. This is a matter of considerable importance, because as Patinkin showed, the operation of the real balance effect in the market experiment is very different from its operation in the individual experiment. We now turn to a discussion of these issues.

III

THE MARKET EXPERIMENT WITH EXOGENOUS NOMINAL MONEY

The way in which a market experiment having to do with the demand for money, the real balance effect, and so on, works out must obviously depend on the nature of the market in which it is performed. It is convenient to begin here with the kind of economy analysed by Patinkin: that is, one in which perfect competition reigns throughout, prices are perfectly flexible, tastes, technology and resource endowments are given and held constant over time, the money supply consists of tokens whose nominal quantity is exogenously given at the beginning of each period, and individual agents face no portfolio adjustment costs. In such a case, one conceivable source of disturbance would be a change in the nominal money supply, and, given perfect price flexibility the outcome of the operation of the real balance effect in the market experiment would be an instantaneous change in the price level. Its effect, as far as the demand for money is concerned would be to keep the economy on its long run function.

This does not imply that it would be econometrically appropriate to substitute the logarithm of actual money supply for m^* on the left hand side of equation (1), and estimate that relationship as a demand for money function, for, with the money supply exogenous and the price level endogenous, these two variables ought to change places. However, because the price level is endogenous in this economy, so are real balances and it would be appropriate to rewrite the long run demand for money function as

$$m^* - p = f(X) \quad (10)$$

Then, provided there was some exogenous variation over time in the factors included in X , a long run demand for real balances function could be estimated in this form if m_s was substituted for m^* .

All this is somewhat academic, since we have noted already that the kind of quarterly and annual data that we use in our empirical work on the demand for money will not permit us successfully to estimate such a long run relationship. We have also noted that it is usual to deal with this problem by adding a lagged dependent variable to the demand function, and that this practise is often defended by referring to the existence of adjustment costs. Suppose that we attempted to introduce these costs into the kind of economy we have briefly described above. Could we account for the presence of a lagged dependent variable in our aggregate demand for money function in these terms? We could not, as I shall now argue. To begin with, recall that equation (3) pictures nominal balances adjusting slowly over time in response to a change in some argument or another of the demand for money function, and note that in the economy I have just described, it is the nominal money supply which changes to disturb agents' money holdings and exogenously at that. In such an economy, where prices are perfectly flexible, individual adjustment costs would have no observable consequences for aggregate behaviour in the face of an exogenous change in the nominal money supply.

The latter assertion seems to fly in the face of certain conclusions that have a well established place in the existing literature of monetary economics. As long ago as (1966) Donald Tucker embedded an aggregate demand for money function essentially the same as equation (3) in an IS-LM model, and showed that the presence of such an equation in a model of that type implies that at least one of the arguments of the demand for money function overshoots its long run equilibrium value as an instantaneous response to a change in the

money supply. This result is mathematically coherent, but it is logically incompatible with the existence of the individual adjustment costs on which the presence of a lagged dependent variable in the demand for money function is usually supposed to be based.

To see why, consider how Tucker's result would apply to an economy in which all the arguments of the demand for real balances (X) are held constant (at their "full employment" levels, perhaps) and in which, therefore, only the price level can adjust to absorb a change in the quantity of money. In such an economy, let the nominal money supply be increased by a certain amount. If there is no lagged dependent variable in the demand for money function that change in the quantity of money will lead to an equi-proportional change in the price level as a result of the pressure of demand exerted on goods markets as all agents try to restore their cash balances. The algebra here is trivial: from (1), if we postulate that the demand and supply of money are to be in equilibrium, we have

$$m_s = m^* = f(X) + p \quad (11)$$

and

$$p = f(X) + m_s \quad (12)$$

so that

$$\frac{\partial p}{\partial m_s} = 1 \quad (13)$$

Now suppose that we maintain the "supply and demand for money are in equilibrium" assumption, but add a lagged dependent variable to the demand for money function. When the nominal money supply changes, a greater than proportional change in the price level seems to be required. From (3) we have

$$p = bf(X) + \frac{1}{b} m_s + \frac{1-b}{b} m_{-1} \quad (14)$$

from which it seems to follow that

$$\frac{\partial p}{\partial m_s} = \frac{1}{b} > 1 \quad (15)$$

This is a very strange result indeed. Faced with portfolio adjustment costs, the individual experiment tells us that the typical agent in the economy is prepared to take time about getting back to equilibrium, and that he therefore changes his demand for goods by less than he would in the absence of such costs when he receives an addition to his holdings of money. Yet we are asked to believe that the aggregate effect of this smaller increase in demand, this weaker real balance effect, is to cause the price level to change by a greater amount than it otherwise would. The conclusion is obvious nonsense.

The problem here has arisen because we have given the wrong interpretation to the variable m_{-1} in the aggregate demand for money function. The individual experiment which must underly the market experiment we are discussing here is one in which the typical agent receives a windfall

gain in money holdings and sets in motion expenditure which enables him to adjust his money holdings towards their long run equilibrium level. As we have seen, in this individual experiment it is crucial to distinguish the cash balances the individual agent chose to end up holding at the end of period -1 on the one hand, and those with which he begins the current period on the other, because these two amounts are not the same when the individual's holdings of nominal money are exogenously disturbed. If he faces portfolio adjustment costs, the individual attempts to move his holdings of nominal money part of the way from where they are at the beginning of the period to the value given by his long run demand for money. That is the meaning of equation (6). The individual can always do this, but when the nominal money supply is exogenous, the whole economy can not. In the aggregate, the money which is available to be held must be held.

But does not equation (15) tell us by how much the price level must change in order for just this to happen? It does not, and the reason why it does not may be seen by considering equation (6), the individual short run demand for money function. The aggregate demand for money is, of course obtained by adding up the latter expression over all individuals in the economy. However, in the experiment we are considering we must substitute the individual's beginning of period money holding for the variable m_{i-1} on the right hand side of (6). Therefore, the current period's money supply rather than the previous period's aggregate demand for nominal money ought to be substituted for the variable labelled m_{-1} on the right hand side of equation (3). If we do this, equation (14) becomes,

$$p = bf(X) + \frac{1}{b} m_s + \frac{1-b}{b} m_s \quad (16)$$

and we have

$$\frac{dp}{dm_s} = 1 \quad (12)$$

Thus, the "overshoot" effect is non-existent, and the economy is always on its long run demand for money function even in the presence of portfolio adjustment costs.

The similar argument can be mounted against Tucker's (1966) analysis, where output and interest rates, rather than prices respond to the change in the money supply and indeed White (1981) essentially does just that. However, let us be clear that what is at stake here is not whether Tucker's results follow from the model that he writes down, because they do, but whether the experiments he carried out with his model are compatible with the underlying adjustment lag assumptions used to justify the presence of a lagged dependent variable in the aggregate demand for money function. An alternative motivation for lags in the demand for money function, for example one based upon sluggishness on the part of expectations to respond to experience, in a world in which the demand for money depends upon the expected rather than actual values of the arguments in the function, would be quite consistent with Tucker's market experiment. His results retain their interest even if we quarrel with one set of premises from which they might be derived. They are in fact very similar (though not identical) to those discussed in Laidler (1968) where the existence of lags in the demand for money function, and elsewhere, is justified, following Friedman (1959), along expectational lines.

The above qualification is of some importance because, a good deal of work on the demand for money function has been devoted to investigating whether or not expectation lags are, in fact, a better explanation than adjustment lags of the need to distinguish between a short run and a long run demand for money function in our empirical work. Specifically, it is well known that, if we substitute the logarithm of real permanent income y^* , and the logarithm of some interest rate r , for X in equation (1), and generate permanent income according to the error learning formula

$$y^* - y_{-1}^* = b(y - y_{-1}^*) \quad (17)$$

Then with the δ s being the parameter of the long run demand for money function, the short run demand for money function is given by

$$m = b\delta_0 + b\delta_1 y + b\delta_2 r + bp + (1 - b)m_{-1} - (1 - b)\delta_2 r_{-1} - (1 - b)p_{-1} \quad (18)$$

If this were the true short run demand for money relationship, we would need to posit no adjustment lags in the individual experiment, agents could always be thought of as holding just the quantity of money they desired, and the long run - short run distinction, upon which successful empirical work seems to depend, would hinge upon discrepancies between current and permanent income.

However, it does not seem possible to defend the proposition that (18) is the true form of the demand for money function in the face of available empirical evidence, even though Feige's (1967) seminal paper on this subject seemed to show that it was. To begin with Feige used annual data, and did

not rule out the possibility that quarterly data might reveal a role for adjustment lags. Subsequent work with quarterly data does seem to show that they have a role to play even in the presence of expectation lags, or at least Laidler and Parkin (1970) claimed that this was the case for the United Kingdom. Goldfeld (1974) went further and concluded that a lag structure of the type captured in equation (3) left nothing for expectations lags to explain in the context of recent United States quarterly data. Furthermore, with similar data, also for the United States, Laidler (1980) found that equation (18) systematically performed worse than an adjustment lag formulation of the function (though not of quite the same form as Goldfeld's as we shall see in a moment). He found that this result held up with annual data too. Moreover, none of this is to mention that recent work on the notion of Rational Expectations must imply that equation (17) is a very dubious formulation of the relationship between permanent income and other observable economic variables.

In short, appealing though the expectations lag hypothesis is as a solution to the problem of linking individual and market experiments while maintaining the distinction between the short and long run demand for money functions, the empirical evidence in favour of this solution is weak. Equations like (3) do fit the data rather well, but if that was because they were really good approximations to equation (18) then the latter would fit even better, and it does not on any systematic basis. However, this does not alter the fact that, in an exogenous money supply world, portfolio adjustment costs cannot be used to motivate the long run - short run demand for money distinction. There is yet a third explanation of the presence of the lagged dependent variable in the aggregate demand for money relationship to be found in the literature on the demand for money. The explanation involves the

"real" (as opposed to "nominal") adjustment model of the short run demand for money. The model is usually written in the following way.

$$m - p = bf(X) + (1 - b)(m_{-1} - p_{-1}) \quad (19)$$

It is then estimated on the assumption that $m_g = m$. By way of comparison, consider equation (3) once more and subtract from both sides of it the logarithm of the current value of the general price level. This yields an expression which differs from equation (19) only in the timing of the value of the price level observation by which lagged nominal balances are deflated.

$$m - p = bf(X) + (1 - b)(m_{-1} - p) \quad (20)$$

Once again, for empirical purposes the money supply is substituted for the quantity of money demanded. Clearly, if one of these expressions fits a particular data set well, so will the other, unless the price level series is extremely erratic, but in practise the series is highly autocorrelated. It was this real adjustment form of the function that Laidler (1980) found to fit better than that derived from an expectations lag and in this study, it also turned out that the real adjustment form performed better than its nominal counterpart in the sense of providing an estimate of $(1 - b)$ that was less than unity. Benjamin Friedman (1977) also obtained this result.

At the level of the individual experiment, this "real" adjustment equation is to say the least decidedly odd. To apply it to the individual experiment is to argue that, if the general price level varies, the typical

agent will instantaneously adjust his nominal balances in order to keep his real money holdings constant, but that a change in any other argument of the long run function will meet with a lagged response. Since the price level is quite as exogenous as any other variable in the individual experiment, and since any adjustment of real balances must involve the agent acquiring or running down nominal balances, it is hard to see why this should be the case. However, when it comes to the market experiment, equation (17) perhaps makes more sense, because it tells us that real balances, rather than exogenous nominal balances, adjust slowly to any disturbance. As we have already noted real balances are endogenous at the level of the market experiment even when nominal balances are not, and Walters suggested as long ago as (1965), that an equation like (19) might be interpreted as a price level adjustment equation in an economy where nominal balances are exogenous.

In fact, equation (19) is not quite accurately specified as a price adjustment equation, as I shall now show. If we have as our aggregate demand for money function

$$m^* = f(X) + p \quad (1)$$

then with an exogenously given money supply, the equilibrium value for the price level p^* is given by

$$p^* = m_s - f(X) \quad (21)$$

Suppose that, for some reason, the structure of the economy was such that the price level moves slowly over time towards equilibrium according to

$$p - p_{-1} = b(p^* - p_{-1}) \quad (22)$$

Then, substitution of (22) into (21), and the addition of m_s to both sides of the equation yields, with a little rearrangement

$$m_s - p = bf(X) + (1 - b)(m_s - p_{-1}) \quad (23)$$

This expression is very like both (19) and (20), and if it was in fact the true relationship describing the way in which the money supply and the arguments of the demand for money function interact over time, one would expect that the other two relationships would display considerable explanatory power as well. Indeed, econometrically speaking, if the "true" relationship was

$$m_s - p = bf(X) + (1 - b)(m_s - p_{-1}) + \epsilon \quad (24)$$

then this would imply, for the "real" adjustment model when m_s is substituted for m

$$m_s - p = bf(X) + (1 - b)(m_{s-1} - p_{-1}) + u \quad (25)$$

where

$$u = \epsilon + (1 - b)\Delta m_s \quad (26)$$

and for the "nominal" adjustment model

$$m_s - p = bf(X) + (1 - b)(m_{s-1} - p) + v \quad (27)$$

where

$$v = \varepsilon + (1 - b) \Delta (m_s - p) \quad (28)$$

The importance of the presence of lagged dependent variables in empirical work on the aggregate demand for money function can then be explained in terms of price level stickiness, and this seems to me to be the best of available explanations. Indeed, it seems to be the only explanation. The portfolio adjustment cost explanation is logically invalid, as we have seen, and the expectations lag explanation does not consistently stand up to empirical testing. However, if this conclusion is accepted, other problems immediately arise, problems which imply that attempts to bring empirical evidence to bear on this issue by attempting to fit equations (24), (25) and (27) to data would be unsatisfactory.

If the price level is sticky, then macroeconomics tells us that such variables as interest rates and real income will tend to change as the money market attempts to clear itself, but these are exactly the variables that one might expect to find in the vector (X). This in turn means that the use of single equation econometric techniques to estimate relationships such as (19), (20) or (23) in an economy where the nominal money supply is believed to be exogenous, is at the very least open to criticism for ignoring simultaneity problems. Jonson (1976) has much to say about this matter in arguing that the demand for money function is best estimated as part of a complete macro system, rather than in isolation, and Cooley and Leroy (1981) have recently raised related issues in discussing the identifiability of the demand for money function.

Econometric questions are undoubtedly important, and it is not my intention to belittle them in any way when I say that, nevertheless, they are not central to the issues that I am attempting to tackle in this paper. The latter are economic in nature. I have shown in the last few pages that adjustment costs at the level of the individual demand for money experiment will produce no observable consequences at the level of the market experiment in an economy in which the nominal money supply is exogenous to the arguments of the demand for money function. It follows from this conclusion that there is something badly wrong with our habit of motivating the long run - short run demand for money distinction in terms of the existence of such adjustment costs. Nevertheless, we seem to need this distinction if we are to deal with a wide variety of real world data on the determinants of money holding a more or less satisfactory way.

I have argued that the best way out of this impasse is to interpret the typical short run demand for money function as a slightly mis-specified price level adjustment equation. If this suggestion is accepted, it follows that the equation which we call a "short run demand for money function" is not a structural relationship at all, but a mixture of structural relationship (the long run demand for money function whose parameters may or may not be being properly estimated if we use single equation techniques), and some reduced form of the whole economy. In particular, the adjustment parameter b must be interpreted as encapsulating the workings of those mechanisms whereby the price level moves slowly towards equilibrium after a monetary disturbance.

The fact that the adjustment speeds which are typically discovered in studies of the aggregate demand for money are very slow - it is seldom

indeed that we find cash balances even half way towards their equilibrium value within a year - has often puzzled monetary economists, but in the light of the above arguments, this fact would appear to tell us that sluggish price adjustment is an important fact of real world economic life. This in turn means that the real balance effect is not just a factor which causes prices to change in some "meta-time" when markets are clearing along Walrasian lines, but is an important empirical phenomenon underlying the generation of real world data, at least in the case where the nominal money supply is exogenous to the arguments of the demand for money function.

IV

A DIGRESSION CONCERNING ENDOGENOUS NOMINAL MONEY

The arguments developed in the last few pages apply to an economy in which the nominal money supply is exogenous to the variables that determine the demand for it. It is sometimes suggested that, in many economies, not least in the United States and the United Kingdom, the actual conduct of policy in recent years has been such as to make it appropriate to think of the nominal money supply as responding passively to demand side factors - in a manner that is reasonably captured by a nominal adjustment version of the short run demand for money function similar to equation (3) or (20). That view is defended in the following way. Whatever changes there may or may not have been in the targets and indicators of monetary policy in the post-war period, its instruments have consistently been interest rates. The monetary authority has attempted to achieve whatever may have been its ends by standing ready to buy and sell securities at a price which although

not necessarily constant over time, is exogenously given at any moment. If over any reasonably short period - say a quarter - real income and prices may be regarded as predetermined, and if the monetary authority, and hence the banking system, stands ready to buy and sell securities at a given price, then there is no obstacle in the way of the economy as a whole adjusting its nominal money holdings towards a desired level at a pace of its own choosing. Given this view of the money supply process, equation (3) is sometimes defended as an appropriate and correctly specified tool for investigating the demand for money.

The argument just presented seems to me to be fallacious. It rests upon a version of what Brunner and Meltzer e.g. (1976) have termed the "money market hypothesis" of the generation of the money supply, adapted to a situation in which the interest rate is the policy instrument, and they have argued that this hypothesis, though widely accepted, is crucially deficient. What Brunner and Meltzer term the "credit market" hypothesis differs from it in insisting correctly that the non-bank public's supply of securities to the banking system is not simply the mirror image of its demand for the liabilities of that system. This is because the non-bank public also holds income earning assets of a type distinct from those that it supplies to the banks when it borrows from them. It is convenient (but not logically necessary) to think of this third asset as reproducible physical capital. To see the significance of this characteristic of the credit market hypothesis for the issues under discussion here it is helpful to begin with a situation of full portfolio equilibrium on the part of the banking system and the non-bank public, and then ask what happens, according to the two hypothesis, when the monetary authorities raise the price at which they are willing to buy securities.

The "money market" hypothesis implies that the public will want to hold more cash balances, and will attempt to acquire them by offering securities to the banking system. It also tells us that any influence on output and prices will come later as a consequence of the effect of the lower interest rate on the level of aggregate demand for goods and services, and that as output and price level changes materialise, more money will be forthcoming from the banking system as the public demands it. In short the nominal stock of money will passively adjust to changes in the arguments of the demand function. If the money market hypothesis is true, nominal money will be just as much an endogenous variable in the market experiment as in the individual experiment, and equation (3) might indeed be an appropriate formulation.

The "credit market" hypothesis leads one to tell a very different story. Certainly a rise in security prices will lead the public to attempt to increase money holdings, but it will also according to this hypothesis, lead agents to attempt to substitute physical capital for securities. In this market experiment, the whole of the non-bank public will try to make such a substitution, and the trick can only be accomplished by selling securities to the banks and taking the proceeds to buy physical capital - but of course the proceeds of such a sale of securities take the form of money, newly created not because the non-bank public as a whole wants to hold it, but because each individual member of that public wants to use money to offer in exchange for capital. Once created, that money must be held, but its creation will coincide with, or even precede, the setting in motion of streams of expenditure that in turn will have consequences for the other arguments of the demand for money function, namely output and prices.

Eventually, the economy will end up with new long run levels of income, prices, money holdings and so on, that may differ little from those which would be predicted by a model that ignored the distinction between securities and physical capital. However, we are here concerned with short run adjustment, with the process whereby this equilibrium is approached, and that is critically different. It involves excess money operating upon expenditure flows that tend to force the arguments of the demand for money function to move towards new values that is to say, it involves real balance effects. This is not to deny that money will also be created and extinguished in such a world in response to changes in the arguments of the demand for money function and in that sense be endogenous, but it is to argue that cause and effect will not run in the simple one-way manner from other variables to money which would be implied by the money market hypothesis, and which would justify equation (3) as a basis for empirical work.

The foregoing arguments fail to touch on yet another reason for not treating the nominal money supply as merely passively responding to the behaviour of the arguments of the demand for money function, even when the monetary authorities treat the interest rate as their principal policy instrument, namely that it is not only the extension of credit to the private sector, but also to the fiscal authorities that leads to the creation of money. Even if, at a particular rate of interest, the values of income, prices, the rate of return on capital, and so on are such as to render the supply of money and bank credit compatible with portfolio equilibrium on the part of the banking system and the non-bank public, that in no way guarantees either that the fiscal authorities' budget is in balance, or that, if it is not, the private sector will be willing to absorb just the right number of

new government bonds to finance whatever deficit is being incurred. A fiscal deficit can therefore become an independent source of monetary expansion when the monetary authorities are treating the interest rate as their policy instrument. Once again cause and effect will run from money creation to variations in the arguments of the demand for money function as well as vice versa and real balance effects will be at work in influencing the outcome of any market experiment.

Closely related to the matters that I have just discussed are considerations having to do with the linkages between the balance of payments and the nominal money supply in an open economy operating a fixed exchange rate. It is true that the balance of payments can provide a channel whereby the nominal money supply will passively adjust to exogenous changes in the arguments in the demand for money function, and that, in a sample of data in which the only, or to put it more practically the major, source of money market disturbance lies in exogenous changes in those arguments, an equation such as (3) might be found to fit the data. In such a case, the parameter b would not be capturing the effects of adjustment costs that face individual agents attempting to re-arrange their portfolios; rather it would be summarising, in one statistic, the structure of the economy's balance of payments mechanism, and in particular the influence of real balance effects on that mechanism. This, however, is only part of the point. In a fixed exchange rate open economy, domestic monetary policy can obviously be an independent source of disturbance to the domestic money market, while any shocks originating in the world economy that have balance of payments side effects can also lead to changes in the money supply that are exogenous to the arguments of the demand function. To say that, in such an economy, the

nominal money supply is endogenous is to say that variation in the nominal money supply is one equilibrating factor at work in the system. To this extent, the monetary behaviour of a fixed exchange rate open economy is different to that of the textbook closed-economy-with-an-exogenous-money-supply model that we considered in the last section of this paper, but the fact of long run endogeneity of the nominal money supply stops far short of establishing the general validity of modelling the short run demand for money along the conventional partial stock adjustment lines embodied in equation (3).

As to the flexible exchange rate case, here we are back to a system which is similar to the closed economy model, at least as far as the relationship between the factors governing the supply of money and the arguments of the demand for money function are concerned. Though the transmission mechanism for the effects of monetary changes to the price level may differ in this case from that to be found in a closed economy, it is nevertheless that transmission mechanism, and not simply the portfolio adjustment costs that face individual agents, that must underly any short-run deviation of actual money holdings and those predicted by the long-run demand for money function.

V

CONCLUSIONS

The arguments presented in this paper have been rather taxonomic, but it is nevertheless possible to draw certain general conclusions from them, conclusions which in their turn yield important implications about

our empirical knowledge of the properties of the demand for money function in particular and of the macroeconomy in general. The basic purpose of this paper has been to argue that the simple portfolio adjustment cost model, on which the distinction between the short run and long run demand for money hinges in the individual experiment, will not do to motivate that same distinction at the level of the economy as a whole. In an economy in which the nominal money supply is exogenous, it is possible for the individual agent to change his holdings of real balances by adjusting his holdings of nominal money, but the whole economy can only accomplish this by changing the general price level. If such an economy is kept "off" its long run demand for money function by adjustment costs, and the data seem to tell us that this is a pervasive phenomenon, then I have argued that the relevant costs are those of changing prices, not those involved in portfolio adjustment.

I have nevertheless argued that, for an economy in which it is believed that the nominal money supply is exogenous to the variables determining the demand for money, various widely used forms of the "short run" demand for money function might deal adequately with the data. They will do so if the complex transmission mechanism that lies between money and prices happens to be such that its dynamics can be captured in that single parameter b , if the simultaneity problems, which must arise here in principle, turn out to be unimportant in practice, and if the money supply and the price level are sufficiently highly autocorrelated that the mis-specifications involved in using the "wrong" lagged dependent variable (c.f., equations (19), (20) and (23)) are also unimportant.

In the case of economies in which the nominal money supply might reasonably be thought of as sometimes adjusting to the arguments of the demand for money function, rather than vice versa, either because of the way in which monetary policy is conducted, or because of the exchange rate regime, the above argument cannot be made as a general proposition. In such economies, whether nominal money does in fact predominantly adjust to demand side factors, or vice versa, depends upon the nature of the shocks to which the economy is being subjected. Even here though, where the shocks in question are such that the nominal money supply is a passively adjusting variable, any slowness on the part of the economy to get back "on" its long run demand for money function, after being driven "off" it, will not simply be a matter of the adjustment costs facing money holders. Instead it will involve the operations of the financial system, and perhaps of the balance of payments mechanism as well.

All this implies that the adjustment parameter b , which in empirical work on the aggregate demand for money plays such an important role in enabling us to get "satisfactory" econometric results, must in general be thought of as summarising the dynamics of a good part of the economic system, and not merely the structural dynamics of the demand for money function itself. In the market experiment, a demand for money function that contains a lagged dependent variable can only be interpreted as being a structural relationship in and of itself if the presence of that variable is justified in terms of expectations lags, and as we have seen, such a justification is hard to sustain as the whole story - though it may be an important component of the story - in the face of available empirical evidence.

The question must arise as to what we are to make of all the empirical evidence that we have on the demand for money function in the light of the foregoing arguments. First, it must not be forgotten that there have been studies of the aggregate demand for money, using long time period samples, and/or data with a high degree of time aggregation that have not invoked the "short run - long run" distinction, and which have generated more or less satisfactory results (e.g., Friedman 1959, Meltzer 1963). Second, and this is important, the vast majority of studies of the "short run" demand for money have produced implicit estimates of the parameters of the "long run" relationship that are reasonably consistent with those derived from the studies just cited. On this basis, I am inclined to argue that, although the interpretation of short-run dynamics that has usually accompanied studies of the short-run demand for money is inappropriate and misleading, nevertheless, for most of the data that have been used, the practise of adding a lagged dependent variable to the function, crude and arbitrary though it is, has turned out in practise to be an adequate way of allowing for the fact that the economy is not always in long run equilibrium when we observe it.

The above argument presents a conjecture, not a well established truth, and it should not lead anyone to conclude that all is well with our empirical knowledge of the demand for money function. In order to find out whether the argument is true, we would have to construct explicit macro-economic models, which permit the economy in general and the money market in particular to deviate from long run equilibrium, and then explicitly investigate the relationship between the structure of those models and the functional forms that have typically been fitted in studies of the short run

demand for money. The parameter b , it has been argued above, is a "black box" parameter that summarises what we may loosely refer to as the dynamics of the real balance effect, and we do not currently know just what is buried in it.

A number of writers, notably Peter Jonson (1976) and Mervin Lewis (1978) have noted that those samples of data that seem to give us the most trouble, as far as finding a stable demand for money function is concerned, are drawn from times and places where the monetary system has been subjected to particularly large shocks, for example the United Kingdom or the United States in the 1970s. They have speculated that the difficulties involved here have stemmed from inadequacy of our modelling of the dynamics of the monetary system. That too is a conjecture, but it is one which follows naturally from the arguments advanced in this paper, and it too could be investigated by carrying out the type of experiment suggested above. The work of Jonson and his associates which builds upon that of Bergstrom and Wymer (1974) has involved the construction of complete econometric models in which expenditure flows of various sorts are explicitly modelled as responding to real balance effects (see e.g., Jonson (1976b), Jonson, Moses and Wymer (1976)) and could provide a framework in terms of which such an investigation could be carried out.

Be that as it may, the arguments that I have advanced imply a severe criticism of much econometric modelling of the conventional post-Keynesian sort. For example, in the FMP model of the United States Economy, a demand for money function complete with lagged dependent variable is estimated independently of the rest of the system, and is treated as a structural relationship to be included in the model, which is then put through all

manner of simulation exercises. If the coefficient of the lagged dependent variable of that demand for money function is in fact capturing, in some approximate and unspecified fashion, aspects of the dynamic behaviour of the economy as a whole, then to treat the relevant function as if it were a structural relationship, and to use it in complete model simulation exercises, is inappropriate. When this is done, the economy's structure is being utilised twice, once in "black box" form in the coefficient of the lagged dependent variable in the demand for money function, and once explicitly in terms of the rest of the model. This criticism, which of course applies to far more pieces of work than the FMP model is in fact a variation on the argument advanced above when we commented on the incompatibility of Tucker's (1966) model with the adjustment lag hypothesis so often used to motivate its structure, and there is no need, therefore, to repeat it here in any detail. Enough has already been said to make the seriousness of its implications for much of our econometric work quite obvious.

The basic conclusion to be drawn from the arguments of this paper is very simple. In treating money as if it were just another durable good in our empirical work on the aggregate demand for money, we have overlooked the critical distinction between the individual and market experiments that Patinkin made so clearly in his theoretical work on the real balance effect. In doing so we have used reasoning that should only be applied to the individual experiment when dealing with the market experiment. As a result, much of our empirical work on the demand for money particularly on the "short run" relationship, has no proper theoretical basis in terms of which it can be interpreted, and our knowledge of that relationship is therefore much less robust than we might have thought. We do, already, have the

basic tools with which we can set about remedying this state of affairs in the shape of those pioneering macro models that explicitly try to get to grips with the dynamics of real balance effects, but the work of developing those models and applying them to the issues raised in this paper has only just begun.

REFERENCES

- Archibald, G. C., and Lipsey, R. G. (1958) "Monetary and Value Theory - A Critique of Lange and Patinkin". Review of Economic Studies, 26, January, 1-22.
- Baumol, W. (1952) "The Transactions Demand for Money - An Inventory Theoretic Approach". Quarterly Journal of Economics, 66, November, 545-556.
- Bergstrom, R., and Wymer, C. R. (1974) "A Model of Disequilibrium Neo-Classical Growth and Its Application to the United Kingdom". London School of Economics International Monetary Research Programme, (mimeo)
- Brunner, K., and Meltzer, A. H. (1976) "An Aggregative Theory for a Closed Economy" in Stein, J. (Ed.), Monetarism, Amsterdam, North Holland.
- Chow, G. (1966) "On the Long-run and Short-run Demand for Money". Journal of Political Economy, 74, April, 111-131.
- Cooley, T., and Leroy, S. (1981) "Identification and Estimation of the Demand for Money". University of California at Santa Barbara, Working Paper 172 (mimeo).
- Feige, E. (1967) "Expectations and Adjustment in the Monetary Sector". American Economic Review, 57, May (papers and proceedings) 462-473.
- Friedman, B. (1977) "The Inefficiency of Short-run Monetary Targets for Monetary Policy". Brookings Paper in Economic Activity, 2, 293-335.
- Friedman, M. (1959) "The Demand for Money - Some Theoretical and Empirical Results". Journal of Political Economy, 67, June, 327-351.

- Friedman, M. (1969) "The Optimum Quantity of Money". Ch. I of The Optimum Quantity of Money, London, Macmillan.
- Goldfeld, S. (1973) "The Demand for Money Revisited". Brookings Papers on Economic Activity, 3, 577-638.
- Harberger, A. C. (Ed.) (1960) The Demand for Durable Goods. Chicago, University of Chicago Press.
- Jonson, P. D. (1976a) "Money Prices and Output - An Integrative Essay". Kredit and Kapital, 9, 499-518.
- _____ (1976b) "Money and Economic Activity in the Open Economy - The United Kingdom 1880 - 1970". Journal of Political Economy, 84, October, 979-1012.
- _____, Moses, E. R., and Wymer, C. R. (1976) "A Minimal Model of the Australian Economy". Reserve Bank of Australia Discussion Paper 7601 (mimeo).
- Keynes, J. M. (1936) The General Theory of Employment, Interest and Money. London, Macmillan.
- Laidler, D. (1976) The Demand for Money - Theories and Evidence (2nd edition), New York, T. Y. Crowell.
- _____ (1968) "The Permanent Income Concept in a Macroeconomic Model". Oxford Economic Papers, 20, March, 11-23.
- _____ (1980) "The Demand for Money in the United States Yet Again" in Brunner, K. and Meltzer, A. H. (Eds.), The State of Macro Economics, Carnegie-Rochester Conference Series, Vol. XII, Amsterdam, North Holland.
- _____, and Parkin, J. M. (1970) "The Demand for Money in the United Kingdom 1956 - 1967: Preliminary Estimates". Manchester School, 38, September, 187-208.

- Lewis, M. (1978) "Interest Rates and Monetary Velocity in Australia and the United States". Economic Record, April, 111-126.
- Meltzer, A. H. (1963) "The Demand for Money - The Evidence from the Time Series". Journal of Political Economy, 71, June, 219-246.
- Modigliani, F. (1977) "The Monetarist Controversy or Should we Forsake Stabilisation Policies?". American Economic Review, 67, March, 1-19.
- Patinkin, D. (1965) Money Interest and Prices (2nd Edition). New York, Harper-Row.
- Tobin, J. (1956) "The Interest Elasticity of Transactions Demand for Cash". Review of Economics and Statistics, 38, August, 241-247.
- _____ (1958) "Liquidity Preference as Behaviour Towards Risk". Review of Economic Studies, 25, February, 65-86.
- Tucker, D. (1966) "Dynamic Income Adjustment to Money Supply Changes". American Economic Review, 56, June, 433-449.
- Walters, A. (1965) "Professor Friedman on the Demand for Money". Journal of Political Economy, 73, October, 545-551.
- White, W. (1978) "Improving the Demand for Money Function in Moderate Inflation". IMP Staff Papers, September, 564-607.